

0923

#6



OIPE

## RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw

3 <110> APPLICANT: BENOIT, Patrick  
 4 SCHWARTZ, Bertrand  
 5 BRANELLEC, Didier  
 6 CHIEN, Kenneth R.  
 8 <120> TITLE OF INVENTION: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING  
 9 THEM AND USES THEREOF  
 11 <130> FILE REFERENCE: 03806.0530-00000  
 13 <140> CURRENT APPLICATION NUMBER: US 10/005,337A  
 15 <141> CURRENT FILING DATE: 2001-12-07  
 17 <150> PRIOR APPLICATION NUMBER: US 60/251,582  
 18 <151> PRIOR FILING DATE: 2000-12-07  
 20 <160> NUMBER OF SEQ ID NOS: 5  
 22 <170> SOFTWARE: PatentIn Ver. 2.1  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 2358  
 26 <212> TYPE: DNA  
 27 <213> ORGANISM: Mus musculus  
 29 <400> SEQUENCE: 1  
 30 ggatcctttc atgtttaaca atatcaaccc taacccaagg ggaacagcct gcctgacagt 60  
 31 ggcttgcca cccatgaata cttcctagtc tagtccgtt gtgaaaactca gcccatttcca 120  
 32 acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttcctg gagcttctt 180  
 33 ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg acccccttgg 240  
 34 gggaatcaaa cgacccttta caggggtcac atatcatcta tcctatatgt caggtattta 300  
 35 cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat 360  
 36 gatttgaaggt caccacaaca tgaggccgccc acactgttct agagaaaaat cacctgggtg 420  
 37 gggaaagggtt tgggaaagcc ttctctgtcca ttcttcattt ttc当地tga tgggttcaca 480  
 38 gaaaggcttt cagctgttct gctggggctc tt当地tgaatc tgatgttggaa ctgtatgtac 540  
 39 caggtctgtc tt当地tgggtt ggagccaaga cgcatcgtgg gtggagcga gacgcaacct 600  
 40 caccttcttag ctctgcattcc atagcaagta gcctaatgtt tctgtgtcta ggtgtcatct 660  
 41 ctgtgaatcg agatccttgg ctttgcatttga attagggagg cacaatatac tcagagattc 720  
 42 aagactgttc agcagcccaag agtcccttcc caaaggaaag gtctcaactc tcagcccccc 780  
 43 ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaaagct gccatagctg 840  
 44 cttgtcactt caagagggtca aagaaaaatag tggtaaccat gaaaacgaga agaccaacag 900  
 45 ttatccattt atagcgttcc aggacagata ggacagagag aacacttagga gagggaacc 960  
 46 cacgaaggac aaggatttag tggatgggtt tt当地ggcaaa tggatgttac tgaagattct 1020  
 47 agaaaacacaa tt当地gttggtaa gaacagctga agtgggggtgg ggggttcttac cccatgttca 1080  
 48 tggagggtt agtggaggaga gacagatata tggatggccag cataacaaac atacacaaca 1140  
 49 ccctaattaa cacttccctc ttctactgac acccccttca ctctccttctt tcataaaaaaa 1200  
 50 taaaaaaagt attttatgtg gcttttacga tagaatctt cctcgaacta taaaaagatc 1260  
 51 taaaatatttta tatttttccat attttatattt cttagcgtt acaagccaga aacaagtattt 1320  
 52 tt当地cccttctt ctcaacagca aagttgggg cttttttttt tccgtgtttag gaatagaaca 1380  
 53 cgagagcccc gtttatcttagt gcaatgttc tatcatttgc ccatgagtct ccagcctcag 1440  
 54 acgcacattt ttctcgggtt ctcttaatgtt tttcccacag cattggggaaa ctttactgac 1500

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/005,337A

DATE: 09/23/2002  
TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt  
Output Set: N:\CRF3\09232002\J005337A.raw

55 agcatccaag ttgtgcttct gctaaagaact ggactcacat ctctctgtgc atcacattcg 1560  
56 cccgtttgg ggtagatcct ctgatttagcc ttcagattta gaacacgggt agcctgtgg 1620  
57 gcactaatta tggccagtga caccatagag tcaaagtgc ttactgaatg ctttcaattt 1680  
58 ctcctaattgc tggtagatg gcatgtcaca gggccatttt agctgcagac atcaactccag 1740  
59 agaattccaa acagatagag acaagtggca cccagaccca ttccttccc ctcgggctga 1800  
60 ttatccccag aaataggatg tcccaaagca acacttccc gccaactggc gtgctgataa 1860  
61 gtccagttat cagaaagata tggctgtaaag tgtatgcac agtgcttgca ttttcttgat 1920  
62 acgttagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta 1980  
63 acaggcagct gtcccctggc ttcccgatac gtgggatgac tcgcattgct gagcgggtgt 2040  
64 gtcactgcca aaggaatgac cctctcacat ttcttcctga ttccgcatacg ccgcggccag 2100  
65 cttgtcatct ccctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct 2160  
66 ctgaattggc cactgggtgg ggcaggggtg tgacttggct tcccaggctg gaagattatc 2220  
67 tcacccagcc ctagctatat aacgggctgg tgtggaggggg ctccacaggg ccagttccag 2280  
68 gggtcatcc acaagagaga aaaacataga ctcgaggctc agggagcttgc catgcctgca 2340  
69 ggtcgaggc caccatgg 2358  
72 <210> SEQ ID NO: 2  
73 <211> LENGTH: 2074  
74 <212> TYPE: DNA  
75 <213> ORGANISM: Homo sapiens  
77 <400> SEQUENCE: 2  
78 ctgcagcaag ttacttaatg tttttgcct cagcatccctc tctgtaaaat gagagcatta 60  
79 gtcttgctcc aacttcgagg gcatggacag ctctggatt tcataatccaa gacccttaaa 120  
80 catcccacag tccttcccccc aaacacttct cctcctaata cctccctcag tttgggtcag 180  
81 gccttggaca aaaaggcata cgaatggta gaaaaagtgt ccatgactac ttctgactta 240  
82 gatgaagaga ccaatgaaaa tagtaatgac tctgttgct tcagcaggac atataactaaa 300  
83 ataggagcta tacaaagaag attagcatgg actctgtgc agaatgacac acaaatttgt 360  
84 gaaacattcc atatattaaa aataaataaaa taataaagag aaaaggaaaa aattaaaaag 420  
85 aaaatagtga tagctgtgtc catctcaaag aaaagcccgag gagatttcct ttatttaccc 480  
86 ccttaagat agaatattag gagaccggaa catatgatac aggaggtact gggagggtcc 540  
87 ctcttgcata atgtttgtc ttgggggtggg gagtcgatgt cttctcaaaat ttcagaaac 600  
88 accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca ttgttgatt 660  
89 gggtagttt ggggagaaat agacacacaa aggtcaaaaca taacttccta attaacactt 720  
90 ccctccattc acaattccct tctccattc ttctctccctg tcttttacts akaraaaccc 780  
91 agttttctt gaaactataa aaatacccccc agtatgtta cataatttac acctcaaaga 840  
92 ttagaaacca gaaatagaga cctttcaac cttccggaa gcaaagtgc ttatccctcc 900  
93 agccacgtgt ctcaaatctt gatgcattcag aatcatctgg gtgcattkaa attcaagatg 960  
94 attctacga gttaccataa atcaactcag aattccctgg agtggggcga gggatctgta 1020  
95 tttctgacaa gctccacacag gtgattcctt tccccacagc attgagaac ttcaactcaa 1080  
96 tgacctaattc agagtccgtc cattgataat atctggtctc atttttbtca tataatata 1140  
97 tagtatttgc ggttagagatg ggattttgcct atgttgccttgc ggcattgtattt gaaactccaa 1200  
98 gctaagcaat cttcctgtct ctgcctccca aatgttggg attacagggtt taagccactg 1260  
99 cacccggctg atagctggtt tcatttactc tatttcttgc ccaactctgat ccattttgaa 1320  
100 gtaaaaaatgc tccaatttattt atgctgtttt agaacacgggt aagcatgtca tgcataatg 1380  
101 gccagtgaca tcataaaaaga aaagtgcattt actgaatgtt ttcaatgtct tataatgtat 1440  
102 gtaagggtggc atgtcatggg gcttatttttgc cccagacatc actccaaaga attccaaaca 1500  
103 gatataagaca agtgccttta gggcccgat ccctccctt caggctgttt acccaggaa 1560  
104 taggatgtcc tgggacaatg ttcccttaag tgaagtgttg ataagtctgc ttatcagaaa 1620  
105 gatattactg ggggtgtgtat atgttagggca tctacattttt cttgatagtt agtcatatga 1680  
106 aagctgacaa agaaaaaaag ggcagtgatg tggtgcaatg tcaacagaca gctgtccct 1740

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/005,337A

DATE: 09/23/2002  
TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt  
Output Set: N:\CRF3\09232002\J005337A.raw

107 gactcttgac aaataggatg acttgcattt ctgagcgatg tgatcaccac caaaggaatg 1800  
 108 gccctctcac atttcttcct gattcacata ttcagcaggg ttagcttgc ctcccctccc 1860  
 109 tcttcagctt cccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc 1920  
 110 taatggggc gggaggttta ctcgggttc caggtggaa gattatctca cccggccca 1980  
 111 gctatataag ctgaccggtg tggagggcc cagcaggcc aactccaggg attccttcca 2040  
 112 cgacagaaaa acataacaaga ctccttcagc caac 2074  
 115 <210> SEQ ID NO: 3  
 116 <211> LENGTH: 750  
 117 <212> TYPE: PRT  
 118 <213> ORGANISM: Homo sapiens  
 120 <400> SEQUENCE: 3  
 121 Met Gly Thr Leu Gly Asp Ser Pro Ile Asp Pro Glu Ser Asp Ser  
 122 1 5 10 15  
 124 Phe Thr Asp Thr Leu Ser Ala Asn Ile Ser Gln Glu Met Thr Met Val  
 125 20 25 30  
 127 Asp Thr Glu Met Pro Phe Trp Pro Thr Asn Phe Gly Ile Ser Ser Val  
 128 35 40 45  
 130 Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile Lys Pro  
 131 50 55 60  
 133 Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr Glu Asp  
 134 65 70 75 80  
 136 Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys Tyr Asp  
 137 85 90 95  
 139 Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro Ala Ser  
 140 100 105 110  
 142 Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro His Glu  
 143 115 120 125  
 145 Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys Gly Asp  
 146 130 135 140  
 148 Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly Cys Lys  
 149 145 150 155 160  
 151 Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp Arg Cys  
 152 165 170 175  
 154 Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys Gln Tyr  
 155 180 185 190  
 157 Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn Ala Ile  
 158 195 200 205  
 160 Arg Phe Gly Arg Met Pro Gln Ala Glu Lys Glu Lys Leu Leu Ala Glu  
 161 210 215 220  
 163 Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp Leu Arg  
 164 225 230 235 240  
 166 Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser Phe Pro Leu  
 167 245 250 255  
 169 Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr Thr Asp Lys  
 170 260 265 270  
 172 Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met Gly Glu Asp  
 173 275 280 285  
 175 Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln Ser Lys Glu  
 176 290 295 300

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/005,337A

DATE: 09/23/2002  
TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt  
Output Set: N:\CRF3\09232002\J005337A.raw

178 Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser Val Glu Ala  
 179 305 310 315 320  
 181 Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly Phe Val Asn  
 182 325 330 335  
 184 Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly Val His Glu  
 185 340 345 350  
 187 Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp Gly Val Leu  
 188 355 360 365  
 190 Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu Lys Ser Leu  
 191 370 375 380  
 193 Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys Phe Glu Phe Ala Val  
 194 385 390 395 400  
 196 Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Leu Ala Ile Phe Ile  
 197 405 410 415  
 199 Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly Leu Leu Asn Val Lys  
 200 420 425 430  
 202 Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln Ala Leu Glu Leu Gln  
 203 435 440 445  
 205 Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu Phe Ala Lys Leu Leu  
 206 450 455 460  
 208 Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr Glu His Val Gln Leu  
 209 465 470 475 480  
 211 Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met Ser Leu His Pro Leu  
 212 485 490 495  
 214 Leu Gln Glu Ile Tyr Lys Asp Leu Tyr Ala Trp Ala Ile Leu Thr Gly  
 215 500 505 510  
 217 Lys Thr Thr Asp Lys Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu  
 218 515 520 525  
 220 Met Met Gly Glu Asp Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln  
 221 530 535 540  
 223 Glu Gln Ser Lys Glu Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe  
 224 545 550 555 560  
 226 Arg Ser Val Glu Ala Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile  
 227 565 570 575  
 229 Pro Gly Phe Val Asn Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys  
 230 580 585 590  
 232 Tyr Gly Val His Glu Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn  
 233 595 600 605  
 235 Lys Asp Gly Val Leu Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu  
 236 610 615 620  
 238 Phe Leu Lys Ser Leu Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys  
 239 625 630 635 640  
 241 Phe Glu Phe Ala Val Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp  
 242 645 650 655  
 244 Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly  
 245 660 665 670  
 247 Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln  
 248 675 680 685  
 250 Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/005,337A

DATE: 09/23/2002

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt  
Output Set: N:\CRF3\09232002\J005337A.raw

251 690 695 700  
253 Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr  
254 705 710 715 720  
256 Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met  
257 725 730 735  
259 Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr  
260 740 745 750  
263 <210> SEQ ID NO: 4  
264 <211> LENGTH: 30  
265 <212> TYPE: DNA  
266 <213> ORGANISM: Artificial sequence  
268 <220> FEATURE:  
269 <223> OTHER INFORMATION: PCR Primer  
271 <400> SEQUENCE: 4  
272 ggcgatttaa ataatgttagt cttatgcaat 30  
275 <210> SEQ ID NO: 5  
276 <211> LENGTH: 31  
277 <212> TYPE: DNA  
278 <213> ORGANISM: Artificial sequence  
280 <220> FEATURE:  
281 <223> OTHER INFORMATION: PCR Primer  
283 <400> SEQUENCE: 5  
284 ggggtctaga aggtgcacac caatgtggtg a 31

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/005,337A

DATE: 09/23/2002

TIME: 10:56:27

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw